

DETEKCIA A SLEDOVANIE ĽUDSKEJ POSTAVY V  
SEKVENCIÍ OBRAZOV

LENKA SLANÁ

2006

DETEKCIA A SLEDOVANIE ĽUDSKEJ POSTAVY V  
SEKVENCII OBRAZOV

DIPLOMOVÁ PRÁCA

LENKA SLANÁ

UNIVERZITA KOMENSKÉHO V BRATISLAVE  
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY  
KATEDRA APLIKOVANEJ INFORMATIKY

počítačová grafika

školiteľ: RNDR. PETER GEJGUŠ

BRATISLAVA 2006

## Čestné prehlásenie

Čestne prehlasujem, že diplomovú prácu som vypracovala samostatne s použitím uvedenej literatúry.

## Podakovanie

Touto cestou, by som chcela vyjadriť moje poďakovanie školiteľovi diplomovej práce RNDr. Petrovi Gejgušovi za čas a všestrannú pomoc, cenné rady a pripomienky pri spracúvaní témy diplomovej práce a testovaní SW diela. Poďakovanie patrí aj dobrovoľníkom, ktorí sú zachytení v testovacích dátach a rodine <sup>1</sup>.

---

<sup>1</sup>najmä mamičke, oteckovi a bračkov

## Abstrakt

V našej práci sme sa zaoberali rozpoznávaním postáv v sekvencii obrazov na základe vlastností kontúry objektu.

Prístup k identifikácii postáv je založený na vlastnostiach vzdialenostnej funkcie, ktorá je v každom obrazovom snímku vyhodnocovaná pre segmentované objekty. Pre  $x$ -ovú súradnicu vzťažného bodu používame aritmetický priemer  $x$ -ových súradníc bodov kontúry.  $Y$ -ovú súradnicu vzťažného bodu sme určili ako 0,8 násobok výšky kontúry.

Na vzdialenostnej funkcii hľadáme lokálne extrémny, ktoré využívame na vytvorenie zjednodušeného modelu kontúry. Tým sme dosiahli veľké zjednodušenie informácie o povahe siluety objektu.

Pre rozpoznanie postavy porovnáваме zjednodušený model kontúry objektu s niekoľkými typmi póz postavy. Deklarovali sme tieto základné typy<sup>2</sup>:

- „I“ (4 extrémny)
- „A“, „I-“ (6 extrémov)
- „T“, „F“, „A-“ (8 extrémov)
- „X“, „E“ (10 extrémov)

Rozpoznanie prebieha v každom snímku video sekvencie, pričom pre každý relevantný objekt prebehne rozpoznanie identity a vyhodnotí sa pomer počtu pozitívne vyhodnotených testov pre daný objekt k počtu všetkých rozpoznávacích testov pre daný objekt.

Tento pomer nám tak v priebehu celej sekvencie poskytuje pravdepodobnostný údaj o identite objektu. V závislosti od vlastností scény, volíme prah, podľa ktorého objekty vyhodnocujeme ako postavy.

---

<sup>2</sup>príčom sme neuvažovali, zriedkavé pózy či odlišnosti spôsobené oblečením

# Obsah

|   |           |
|---|-----------|
| Čestné prehlásenie . . . . .                          | i         |
| Poďakovanie . . . . .                                 | ii        |
| Abstrakt . . . . .                                    | iii       |
| <b>1 Úvod</b>   | <b>1</b>  |
| <b>2 Segmentácia</b>                                  | <b>4</b>  |
| 2.1 Adaptívne pozadie (Adaptive Background) . . . . . | 6         |
| 2.1.1 Výber metódy . . . . .                          | 6         |
| 2.1.2 Činnosť metódy . . . . .                        | 7         |
| 2.2 Filtrácia tieňov . . . . .                        | 9         |
| <b>3 Sledovanie postavy</b>                           | <b>12</b> |
| 3.1 Výber metódy . . . . .                            | 12        |
| 3.2 Pravdepodobnostné metódy . . . . .                | 13        |
| 3.3 Kalmanov Filter . . . . .                         | 13        |
| 3.3.1 Činnosť metódy . . . . .                        | 14        |
| 3.4 Časticový Filter (Particle Filter) . . . . .      | 15        |
| 3.4.1 Činnosť metódy . . . . .                        | 16        |
| <b>4 Špecifikácia softvérového diela</b>              | <b>18</b> |
| 4.1 Úvod . . . . .                                    | 18        |

|          |  |           |
|----------|--|-----------|
| 4.2      | Vstupné dáta . . . . .                                     | 18        |
| 4.2.1    | Kamera . . . . .   | 18        |
| 4.2.2    | Postavy . . . . .  | 19        |
| 4.3      | Výstupné dáta . . . . .                                    | 19        |
| <b>5</b> | <b>Implementácia</b>                                       | <b>20</b> |
| 5.1      | DirectShow . . . . .                                       | 20        |
| 5.1.1    | GraphEdit . . . . .  | 21        |
| 5.2      | Parametre nastaviteľné užívateľom . . . . .                | 23        |
| 5.2.1    | Adaptive Background . . . . .                              | 24        |
| 5.2.2    | Shadow . . . . .   | 27        |
| 5.2.3    | Blob . . . . .   | 28        |
| 5.2.4    | Contour . . . . .  | 30        |
| 5.2.5    | Tracking . . . . .   | 32        |
| <b>6</b> | <b>Metóda na rozpoznávanie postáv (pedestrian tracker)</b> | <b>33</b> |
| 6.1      | Relevantné práce . . . . .                                 | 33        |
| 6.2      | Segmentácia . . . . .                                      | 34        |
| 6.2.1    | Dodatočné spracovanie . . . . .                            | 35        |
| 6.2.2    | Zrýchlenie . . . . .                                       | 36        |
| 6.3      | Analýza masky – hľadanie objektov . . . . .                | 37        |
| 6.4      | Rozpoznanie postavy . . . . .                              | 38        |
| 6.4.1    | Kostra . . . . .   | 39        |
| 6.4.2    | Vzdialenostná funkcia . . . . .                            | 40        |
| 6.4.3    | Kostra alebo vzdialenostná funkcia ? . . . . .             | 40        |
| 6.4.4    | Hľadanie extrémov . . . . .                                | 41        |
| 6.4.5    | Vzťažný bod vzdialenostnej funkcie . . . . .               | 43        |
| 6.4.6    | Rozpoznanie postavy . . . . .                              | 44        |

|          |                                   |           |
|----------|-----------------------------------|-----------|
| 6.4.7    | Rozpoznanie postavy . . . . .     | 47        |
| 6.4.8    | Presnosť rozpoznania . . . . .    | 48        |
| <b>7</b> | <b>Testovacie dáta a výsledky</b> | <b>50</b> |
| 7.0.9    | Testovací stroj . . . . .         | 51        |
| 7.1      | Rýchlosť . . . . .                | 51        |
| 7.2      | Presnosť . . . . .                | 52        |
| 7.2.1    | Exteriér . . . . .                | 53        |
| 7.2.2    | Interiér . . . . .                | 53        |
| 7.2.3    | Syntetické scény . . . . .        | 53        |
| 7.3      | Výsledky iných . . . . .          | 54        |
| <b>8</b> | <b>Budúca práca</b>               | <b>55</b> |
| <b>9</b> | <b>Záver</b>                      | <b>56</b> |



# Kapitola 1

## Úvod

V tejto diplomovej práci sa zaoberáme zaujímavým problémom rozpoznávania postáv v sekvencii obrazov. Otázkou prirodzenou pre ľudské oko a myseľ, nie však samozrejmom pre kameru a počítač.

Stroje využívame na to, aby za nás pracovali, počítače na to aby za nás rátali a snímacie zariadenia, aby za nás vnímali. A predsa to nestačí, pretože ľudia vedia viac ako pracovať, rátať a vnímať. Ľudia rozumejú svetu v ktorom sa nachádzajú, alebo sa o to aspoň snažia a ak dochádzajú na okraj svojich možností hľadajú niekoho – niečo koho možnosti siahajú ďalej. Preto naučiť stroje rozumieť veciam a javom okolo seba je mýtom a zároveň cieľom vedy, presne tak ako prekonanie vlastných možností a schopností či túžba lietať.

Najzákladnejším ľudským vnemom je zrak. Analógiou pre počítače sa zaoberá vedná disciplína počítačové videnie a rozpoznávanie postavy zo sekvencie obrazov, je jeho neodmysliteľnou súčasťou.

Treba si však hneď uvedomiť, že človek má na riešenie tej istej otázky množstvo skúseností, ktoré do jeho života prichádzajú po narodení a prichádzajú relatívne pomaly a s relatívne dokonalou technickou výbavou.

Problém rozpoznávania postavy je možné deliť do menších skupín podľa kritérií pre kameru, objekty v scéne a scénu nasledovne.

Podľa typu, počtu a umiestnenia kamier:

- optické / termo kamery
- 1 kamera – mono videnie / 2 kamery – stereo videnie, rozpoznávanie po rekonštrukcii 3D modelu
- statické / pohyblivé kamery

Podľa typu a počtu a pohybu objektov v scéne:

- osoby, dopravné prostriedky, zver, predmety ...
- prekrývajúce sa, zmenšujúce sa ...

Podľa typu scény:

- štúdiové rozpoznávanie (špeciálne upravená scéna)
- interiér
- exteriér

V našej štúdii sme sa zamerali na vstup z jednej kamery, na typ scény sme nekládli žiadne obmedzenia, ale pre pozitívne rozpoznané postavy sme určili nasledovné podmienky: vzpriamené držanie tela, neprekrývanie postáv a vylúčili sme postavy s predmetmi. Pre tieto vstupné podmienky sa dá proces



Obr. 1.1: Tri základné kroky procesu rozpoznávania.

rozpoznania postavy, rozdeliť na tri základné kroky, znázornené na obrázku 1.1, ktoré sú vykonávané pre každý snímok a ktorých význam presnejšie priblížime v kapitolách 2, 3, 6.

Pomocou kamery získavame obrazové snímky, ktoré využívame ako vstupné údaje pre segmentáciu. Segmentácia slúži na vyčlenenie objektov zo snímku ako celku a to na základe rozlíšenia pozadia od popredia scény (kapitola 2). Sledovanie (tracking) objektu slúži na hľadanie vhodnej polohy postavy a pozorovanie pohybových zmien objektu (kapitola 3).

Pri rozpoznávaní postáv nesmieme zabudnúť na to, že nie každá situácia je jednoznačne rozpoznateľná. Toto je otázka relevantná hlavne pre uplatnenie metód rozpoznania v oblasti bezpečnostnej a strážnej techniky. Treba však pripomenúť, že úroveň rozpoznávania postáv nie je tak pokročilá aby dosahovala kvality ľudských schopností. Všetci dobre vieme, že pri rozpoznávaní postáv môže zlyhať aj ľudský faktor, keďže postava ktorá má nečistý úmysel sa môže veľmi dobre maskovať.

# Kapitola 2

## Segmentácia

Nevyhnutnou časťou rozpoznávania – identifikácie objektov je ich segmentácia, ktorá rozdeľuje obraz na objekty a tým nám poskytuje prvotnú informáciu o možnom výskyte postavy. Je to základná časť celého rozpoznávania postavy, pretože na kvalite segmentácie závisí aj kvalita výsledku rozpoznania.

Výsledkom segmentácie býva binárna maska, odlišujúca pohyblivé objekty od pozadia, ktorá reprezentuje vstup pre ďalšie spracovanie.

Spôsoby segmentácie sa rôznia v povahe vstupných a výstupných dát, preto sa stretávame s rôznymi typmi segmentácie, ktoré vo svojej práci na základe segmentačnej metódy kategorizovali a opísali [1] nasledovne:

### 1. pohybovo zameraná segmentácia (Motion-based)

#### A) 2D:

##### a) optické toky (Optical flow):

i) štatistické metódy

ii) hľadanie blokov (Block matching)

ii) diferenčné metódy (Differential method)

- b) detekcia zmeny (Change Detection):
  - i) rozdiely v akumulovaných snímkoch
  - ii) rozdiely v snímkoch

B) 3D:

- a) parametrické metódy (Parametric):
  - i) 6 parametrov
  - ii) 8 parametrov
  - iii) 12 parametrov
- b) SFM
  - i) nelineárne
  - ii) lineárne

2. časo-priestorová segmentácia (Spatio-temporal)

A) časové (Temporal):

- a) parametrické
- b) sledovanie zmien
- c) optické toky

B) priestorové (Spatial):

- a) zamerané na oblasť (Region-based):
  - i) rast regiónov (Region growing)
  - ii) Bayesian
  - iii) Watershed
- b) zamerané na kontúru (Contour-based):
  - i) Sobelov hranový detektor (Sobel edge detector)
  - ii) Cannyho hranový detektor (Canny edge detector)

S ohľadom na naše vstupné podmienky sme sa pri výbere metódy segmentovania zamerali na metódy pracujúce na základe dočasných zmien (v 2 – 3 posledných snímkov, kategória 1.A.b. ) [1], [2], optický tok [1], [2] [3] (kategória 1.A.a) a odhad pozadia [3], [4], [5], [6], (kategória 2.A.b).

Segmentácia podľa *dočasných zmien*, je veľmi vhodná pre rýchlo sa meniace scény, ale vo všeobecných prípadoch nezachytáva všetky relevantné obrazové body. *Optické toky* môžu byť použité pri selekcii nezávisle pohybujúcich sa objektov pri použití pohyblivo umiestnenej kamery. Sú výpočtovo veľmi náročné a v súčasnosti nepoužiteľné pre aplikácie určené pre nešpecializovaný hardvér s činnosťou v reálnom čase, preto sme sa rozhodli pre *odhad pozadia* metódou adaptívneho pozadia (adaptive background method).

## 2.1 Adaptívne pozadie (Adaptive Background)

### 2.1.1 Výber metódy

Keďže náš celkový cieľ je vytvoriť robustný systém pracujúci v reálnom čase, rozhodli sme sa využiť na segmentáciu metódu adaptívneho pozadia (*adaptive background method*), ktorá vytvára model pozadia na základe skupiny niekoľkých gaussovských pravdepodobnostných rozložení.

Tento prístup umožňuje korektne pracovať či už v interiéroch alebo v exteriéroch, nezachytáva pohyby listov svetelné zmeny na vodnej hladine, jasové zmeny pozadia spôsobené zmenou svetelnosti (napr. pohyb mrakov) či záchvevy kamery, [3], [4], [5], [6].

## 2.1.2 Činnosť metódy

Každé gaussovské pravdepodobnostné rozloženie z použitej skupiny, reprezentuje hodnotu pre pixel pozadia. V prípade, že sa hodnota pixla od stredných hodnôt gaussovských pravdepodobnostných rozložení dostatočne odlišuje, je daný pixel vyhodnocovaný ako popredie a to až do vtedy, kým nevyhovuje niektorému zo skupiny gaussovských pravdepodobnostných rozložení.

V ľubovoľnom čase si pre každý pixel  $(x_0, y_0)$  pamätáme jeho históriu, čiže informáciu o farbe v podobe „RGB“ vektora, pričom  $I$  predstavuje danú obrazovú sekvenciu.

$$\{X_1, \dots, X_t\} = \{I(x_0, y_0, i) : 1 \leq i \leq t\}$$

História každého pixlu sa modeluje pomocou  $K$  gaussovských pravdepodobnostných rozložení. Pomocou nich ďalej vyjadríme pravdepodobnú hodnotu pixla  $P(X_t)$ . Kde  $K$  je počet distribúcií,  $\omega_{i,t}$  váha odhadu  $i$ -teho gausiánu v čase  $t$  (vyjadruje aká časť dát má byť zarátaná do gausiánu),  $\mu_{i,t}$  predstavuje hodnotu  $i$ -teho Gausiánu v čase  $t$ ,  $\Sigma_{i,t}$  je kovariančná matica  $i$ -teho gausiánu v čase  $t$  a  $\eta$  je funkcia hustoty pravdepodobnosti gausiánu.

$$P(X_t) = \sum_{i=1}^K \omega_{i,t} * \eta \left( X_t, \mu_{i,t}, \Sigma_{i,t} \right)$$
$$\eta \left( X_t, \mu, \Sigma \right) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(X_t - \mu)^T \Sigma^{-1} (X_t - \mu)}$$

$K$  sa určí podľa možností pamäte a výpočtovej sily stroja. V súčasnej dobe sa používa v rozmedzí 3 až 5, pričom na činnosti metódy sa voľba vyššieho  $K$  prejaví v prípadoch, keď bod pozadia nadobúda  $K$  možných podôb. Z výpočtových dôvodov predpokladáme že kovariančná matica má nasledujúcu formu:

$$\sum_{k,t} = \sigma_k^2 I$$

Každý nový pixel,  $X_t$ , sa prehodnotí  $K$  gaussovskými pravdepodobnostnými rozloženími, až kým niektoré z nich nevyhovuje. Vyhovovať znamená, že hodnota pixla je v rozpätí 2,5 smerodajnej odchýlky distribúcie. Je to efektívny prah typu pixel na distribúciu.

Ak nový pixel nevyhovuje žiadnej distribúcii, nahradí sa najmenej pravdepodobné gaussové rozloženie novým, pričom ako stredná hodnota gausiánu sa použije hodnota nového pixla inicializovaná na vysokú varianciu a nízku váhu priority.

Váha priority  $K$ -tej distribúcie v čase  $t$ ,  $\omega_{k,t}$  je vyjadrená nasledovne:

$$\omega_{k,t} = (1 - \alpha)\omega_{k,t-1} + \alpha(M_{k,t})$$

kde  $\alpha$  je koeficient učenia a  $M_{k,t}$  je 1 pre model ktorý vyhovoval a 0 pre ostatné modely. Po aproximácii, sa váhy normalizujú.  $1/\alpha$  definuje časovú konštantu ktorá určí akou rýchlosťou sa parametre v distribúcií menia.  $\mu$  a  $\sigma$  sú pre nevyhovujúce distribúcie nezmenené. Pre vyhovujúce distribúcie sa aktualizujú nasledovne:

$$\mu_t = (1 - \rho)\mu_{t-1} + \rho X_t$$

Kde

$$\rho = \alpha\eta(X_t|\mu_k, \sigma_k)$$

Jednou z najväčších výhod tejto metódy je, že v prípade ak na scénu príde objekt ktorý v nej zotrva tak dlho až splynie s pozadím – prestane byť



vyhodnocovaný ako popredie, nenaruší informáciu o orginálnom pozadí, pretože tá je obsiahnutá v zvyšných gaussovských pravdepobnostných rozložení aj keď v tejto situácii jej prislúcha nižšia váha. Takže ak objekt opustí svoju pozíciu orginálne pozadie rýchlo naberie na váhe a metóda opäť poskytuje správne výsledky.

V závere taktu sa po obnovení všetkých parametrov vykoná utriedenie gaussovských pravdepobnostných rozložení podľa  $\omega/\sigma$  od vyhovujúcich po najpravdepodobnejšie distribúcie pozadia. (pole gaussovských pravdepobnostných rozložení utriedime od najpravdepodobnejších – s 0–tým indexom v poli po najmenej pravdepodobné). Za pozadie potom vyberie prvých  $B$  distribúcií, pričom  $B$  volíme podľa rovnice (2.1).

$$B = \underset{b}{\operatorname{argmin}} \left( \sum_{k=1}^b \omega_k > T \right) \quad (2.1)$$

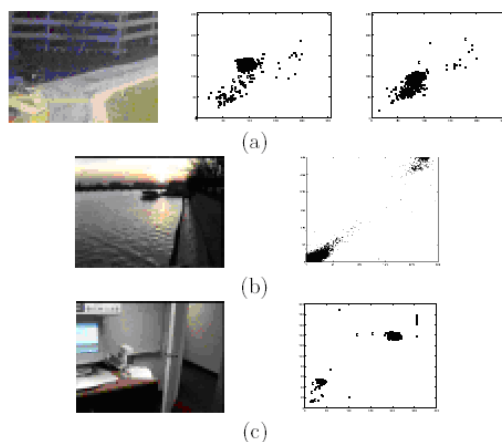
kde  $T$  je miera minimálneho podielu dát, ktoré môžu byť považované za pozadie. Ak je hodnota  $T$  vysoká, dostávame multimodálny vzor pozadia, ktorý spôsobí že opakujúce sa zmeny (napr. pohyb listov na strome) budú vyhodnotené ako pozadie. Voľba nízkeho  $T$ , spôsobí opak.

## 2.2 Filtrácia tieňov

Jedným z najčastejších problémov počítačového videnia pri spracúvaní externých alebo interiérových scén sú tieňe. Pri určovaní pozadia sa tieňe vyhodnotia ako súčasť pohyblivého objektu, čo je pre náš systém nežiaduci jav.

Problém detekcie a filtrácie tieňov je veľmi náročný na robustnosť. V literatúre sa stretávame s niekoľkými prístupmi [7], [8], [9] a [10].

Pre jednoduchosť a intuitívnosť sme sa rozhodli pre metódu založenú na



Obr. 2.1: Práca metódy adaptívneho pozadia znázornená (scatter plot) diagramom pre zelené a červené hodnoty daného obrazového bodu. (a) – dva zhluky toho istého pixlu počas dvoch minút. (b) – bi-modálna distribúcia povrchu vody. (c) – bi-modálna distribúcia monitoru (blikania)

jasových vlastnostiach obrazu a na zmenách v jeho farebnej sýtosti [10], [11].

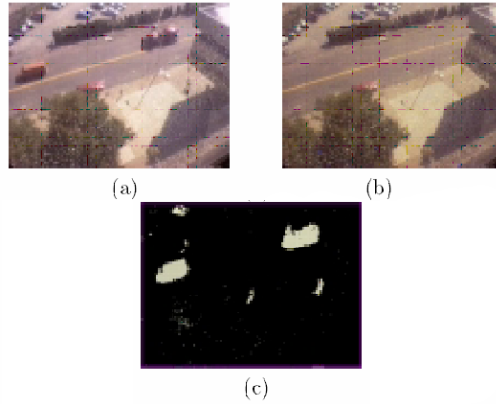
Metóda pracuje s HSV (Hue Saturation Value) farebným modelom v ktorom je navrhnutá klasifikácia tieňov. Idea spočíva v odhadnutí súvislosti medzi identifikáciou tieňa a H, S, V hodnotami obrazu. Konkrétne vychádza z predpokladu, že prítomnosť tieňa spôsobí, že príslušné pixle sú tmavšie a menej farebné.

Z prvého predpokladu o zmene vlastností pixla ktorý je v tieni dostávame vzťah:

$$\frac{I_i^V}{B_i^V}$$

kde  $I$  predstavuje aktuálny snímok,  $B$  reprezentuje model pozadia, index  $i$  určuje  $i$ -ty pixel,  $V$  naznačuje V hodnotu z modelu HSV. V oblasti tieňa dostávame z daného vzťahu vysoké hodnoty a dobre definovaný rozsah.

Druhý predpoklad nám taktiež pomáha rozlíšiť bod pozadia od zatie-



Obr. 2.2: Práca metódy adaptívneho pozadia. (a) – pôvodný obraz. (b) – obraz vytvorený na základe najpravdepodobnejších gaussovských pravdepodobnostných rozložení modely adaptívneho pozadia. (c) – pixle popredia

neného bodu. Analogickým vzťahom, teraz pre hodnoty S z modelu HSV dostávame hodnoty s dobre definovaným rozsahom.

Môžeme teda definovať triedu  $SP_k$  (Shadow point mask) nasledovne:

$$SP_k = \begin{cases} 1 & \text{if } \alpha \leq \frac{I_k^V(x,y)}{B_k^V(x,y)} \leq \beta \\ & \wedge I_k^S(x,y) - B_k^S(x,y) \leq \tau_S \\ & \wedge |I_k^H(x,y) - B_k^H(x,y)| \leq \tau_H \\ 0 & \text{otherwise} \end{cases}$$

kde  $I_k(x,y)$  a  $B_k(x,y)$  predstavuje farebnú hodnotu pixla  $(x,y)$  v danom snímku  $k$  a modeli naučeného pozadia pre snímok  $k$ . Použitie  $\alpha$  dovoľuje vyhnúť sa identifikácii bodov ako tieňa, kde bolo pozadie ovplyvnené šumom, nakoľko  $\beta$  predstavuje „silu“ tieňa, čo je intenzita podľa podmienok prostredia.

Kanál S prahujeme podľa rozdielu. Prahovanie kanála H podľa absolútneho rozdielu dáva lepšie výsledky.

# Kapitola 3

## Sledovanie postavy

Sledovanie postavy môžeme zredukovať na problém sledovania krivky v obraze. Sledovanie je najbežnejšie používané na odstránenie rušenia po segmentovaní. Rušením rozumieme, napríklad ak pozadie obsahuje aj objekty popredia (tento prípad nastáva pri sledovaní jednej postavy prechádzajúcej davom).

### 3.1 Výber metódy

Keďže naše vstupné podmienky nezahŕňajú prípady, v ktorých môže dôjsť k prekrytiu jedného objektu druhým (z pohľadu kamery), uplatňujeme predikciu sledovania iba na urýchlenie činnosti systému (kapitola 6.2.2). Pre tieto účely sme zvolili sledovanie Kalmanovým [2], [4], [12], [13] a časticovým filtrom (Particle – ConDensation filter) [2], [4], [14],[15], ktoré sú založené na pravdepodobnostných metódach.

## 3.2 Pravdepodobnostné metódy

Tieto metódy vyžadujú sledovanie systému situácií, ktoré sa menia v čase, tieto situácie nazývame stavmi a často sú vyjadrované vektorom.

Model predstavuje dátovú štruktúru, ktorá reprezentuje reálne informácie z vizuálnej scény.

Na odhad stavu systému predpokladáme vhodnú a presnú znalosť modelu a jeho parametrov. Parametre predstavujú veličiny, ktoré popisujú model, ale menia sa pomalšie ako stav. Často sú známe a statické. Vektor merania, reprezentuje vstupné dáta zo senzora.

Na reprezentáciu modelu je špecifikované:

- odhadovaná dynamika zmeny stavu z jedného časového momentu do druhého
- metóda na získanie vektora merania.

## 3.3 Kalmanov Filter

Vo všeobecnosti nekladie požiadavku stacionarity a preto môže poskytovať optimálne odhady aj pre nestacionárne signály. Okrem toho neobmedzuje apriori rozsah zberu informácie o signále na posledných  $N$  vzoriek, ale využíva všetky dostupné empirické informácie, ktoré poskytuje pozorovaný signál od počiatku merania či príjmu.

Štruktúra Kalmanovho Filtra je rekurzívna, pričom jeho koeficienty sa v každom takte upravujú na základe dostupnej informácie tak, aby poskytli optimálny odhad signálu.

Zavedenie uvedeného typu modelu umožnilo počítať premenné parametre filtra rekurzívne, t.j. nový filter v každom kroku vzniká opravou filtra z pred-

chádzajúceho kroku na základe získanej novej informácie, bez toho, aby bolo potrebné uchovávať všetky predchádzajúce hodnoty vstupného signálu.

### 3.3.1 Činnosť metódy

Predpokladom pre jednorozmernú Kalmanovu filtráciu je, že signál môže byť modelovaný pomocou vektorového dynamického modelu prvého rádu. Tento vektor sa nazýva stavový vektor. Stavový vektor zahrňuje v sebe minimálne množstvo informácie o predchádzajúcich odhadoch ako aj informácie budúcej odpovede, danej budúcou zašumenou vzorkou na vstupe systému.

Daný dynamický model je opísaný nasledujúcimi rovnicami:

$$x(m) = Fx(m-1) + Gw(m)$$

$$y(m) = Hx(m) + v(m)$$

kde  $x(m)$ ,  $x(m-1)$  sú signálové stavové vektory,  $y(m)$  je pozorovaný signál,  $w(m)$  je vektor šumu náhodného procesu,  $v(m)$  je pozorovaný štatistický šum,  $F$  je systémová matica,  $G$  je budúca matica a  $H$  je sledovacia matica. V rovniciach sú  $w(m)$  a  $v(m)$  dva navzájom nekorelované Gaussove biele šumy, ktoré majú nulovú strednú hodnotu a koreláciu danú:

$$\begin{aligned} E [w(m)w^T(n)] &= Q_w \delta_{mn} \\ E [v(m)v^T(n)] &= Q_v \delta_{mn} \end{aligned} \quad (3.1)$$

kde  $Q_w$  a  $Q_v$  sú korelačné matice a  $\delta_{mn}$  je Kroneckerova delta, pre ktorú platí:

$$\delta_{mn} = \begin{cases} 1 & \text{pre } m = n \\ 0 & \text{pre } m \neq n \end{cases} \quad (3.2)$$

Algoritmus na výpočet Kalmanovej filtrácie pre tento uvažovaný model je potom nasledovný:

$$\hat{x}_b(m) = F\hat{x}_a(m-1) \quad (3.3)$$

$$P_b(m) = FP_a(m-1)F^T + GO_wG^T \quad (3.4)$$

$$K(m) = P_b(m)H^T(HP_b(m)H^T + Q_v)^{-1} \quad (3.5)$$

$$\hat{x}_a(m) = \hat{x}_b(m) + K(m)(y(m) - H\hat{x}_b(m))$$

$$P_a(m) = (I - K(m)H)P_b(m)$$

pričom

$$P_a(m) = E[(x(m) - \hat{x}_a(m))(x(m) - \hat{x}_a(m))^T]$$

$$P_b(m) = E[(x(m) - \hat{x}_b(m))(x(m) - \hat{x}_b(m))^T]$$

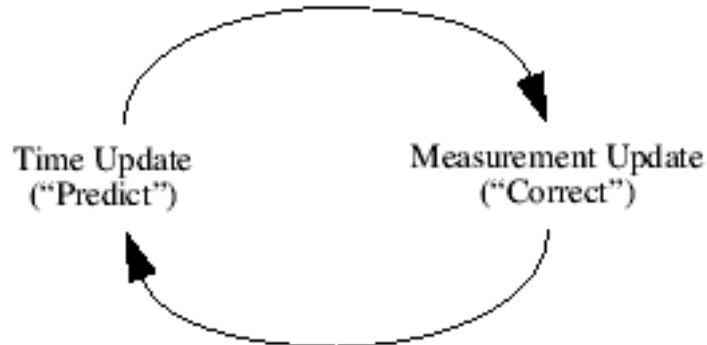
kde  $P_a(m)$  je kovariančná matica vektora filtračnej chyby,  $P_b(m)$  je kovariančná matica vektora predikčnej chyby,  $x(m)$  je vektor skutočných hodnôt,  $\hat{x}(m)$  je estimovaná hodnota. Indexy  $a, b$  vyplývajú zo slov „after“ a „before“.

Všetky uvedené rovnice majú jednoduchú interpretáciu. Rovnica (3.1) popisuje poslednú predikciu, využívajúcu dynamiku systémového modelu. Potom v rovnici (3.4) dochádza k vylepšeniu tejto predikcie využívajúc nové pozorovanie  $y(m)$ . Rovnica (3.3) slúži na získanie matice zosilnenia pre vylepšenie odhadu v rovnici (3.4). Zvyšnými rovnicami (3.2) a (3.5) sa vypočítajú kovariančné matice chyby potrebné pre výpočet matice zosilnenia.

### 3.4 Časticový Filter (Particle Filter)

S časticovým filtrom sa v literatúre stretávame aj pod názvom ConDensation (conditional density propagation) Algorithm.

Je založený na štatistickej metóde Monte-Carlo, takže sa vyhýba analytickým výpočtom. Idea algoritmu sa opiera o použitie náhodne vygenerovaných



Obr. 3.1: Kalmanov filter - schéma činnosti

vzoriek na aproximáciu pravdepodobnostnej funkcie. Vzorky sú náhodne generované stavové vektory a pri štarte metódy ich inicializovať manuálne.

### 3.4.1 Činnosť metódy

Metóda pracuje na lineárnom stochastickom dynamickom systéme:

$$x_{k+1} = Ax_k + w_k \quad (3.6)$$

S meraním  $z$ .

Počas prvej fázy algoritmu sa každá vzorka aktualizuje podľa rovnice (3.6).

Skôr ako získame vektor miery  $z$ , metóda odhadne podmienené pravdepodobnosti hustôt každej vzorky  $p(X_n|z)$  na základe Bayesovho pravidla z rovnice (3.7).

$$p(x|z) = kp(z|x)p(x) \quad (3.7)$$

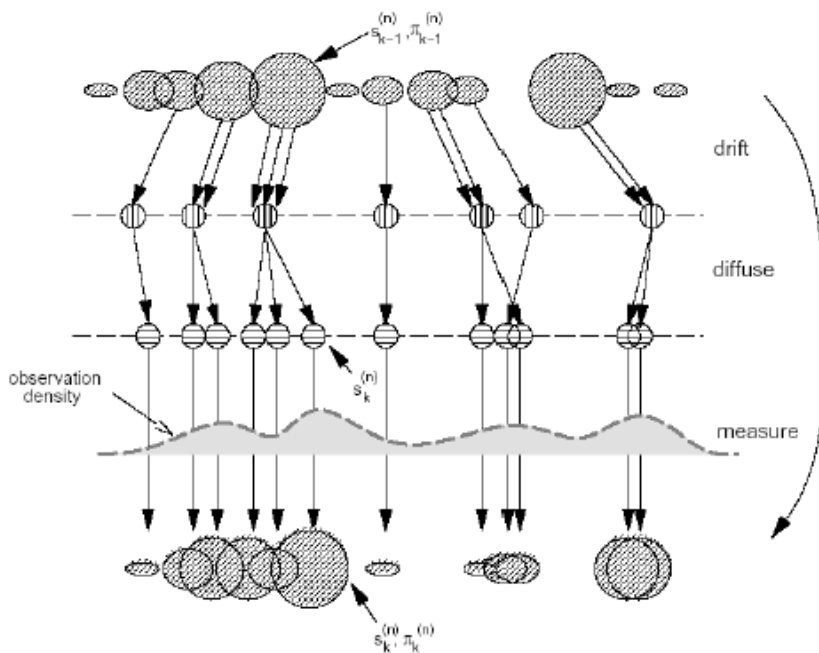
kde  $k$  je normalizačná konštanta nezávislá na  $x$ .



Po vypočítaní pravdepodobností možno vyhodnotiť napríklad význam sledovaného procesu v konkrétnom čase (obr. 3.2).

Vo všeobecnosti možno metódu opísať v troch základných krokoch:

1. **Výber**  $n$ -tej vzorky v čase  $t$ , zo starej  $j$ -tej vzorky pre čas  $t - 1$  prerátaním pravdepodobnosti pre čas  $t - 1$  a vzorku  $j$ . Čo sa dá účinne spraviť použitím kumulatívnych váh (konštruovaných v treťom kroku).
2. **Predpovedať** náhodným vzorkovaním z podmienených hustôt pre dynamický model na vytvorenie vzorky pre novú množinu vzoriek.
3. **Miera** za účelom generovania váh pre nové vzorky. Každá váha je vyhodnotená z „pozorovanej“ hustoty pravdepodobnosti, ktorá je vo všeobecnosti obyčajne multimodálna.



Obr. 3.2: Časticový filter - schéma činnosti

# Kapitola 4

## Špecifikácia softvérového diela

### 4.1 Úvod

Cieľom našej aplikácie je zo vstupnej video sekvencie rozpoznať a vizuálne znázorniť na výstupe nášho systému pozície objektov, ktoré reprezentujú postavu.

### 4.2 Vstupné dáta

Keďže celková problematika rozpoznávania postáv je nad rámec diplomovej práce, vybrali sme si pre našu metódu nasledujúce obmedzenia na charakter vstupnej sekvencie.

#### 4.2.1 Kamera

Scénu sníma statická kamera, ktorá poskytuje farebnú (RGB) video sekvenciu. Táto kamera sníma scénu pod uhlom  $0^\circ$  až  $60^\circ$ .

### 4.2.2 Postavy

Za kladne rozpoznané budeme vyhodnocovať najmä postavy, ktoré:

- sú v scéne viditeľné celé (žiadna časť ich tela nie je mimo obrazový snímok)
- nemajú predmety (uchopené v rukách, alebo inak spojené s telom)
- majú vzpriamené držanie tela (hlava je nad úrovňou brušnej časti, pričom postava nie je v podrepe, alebo ležiaca) a neprekrývajú sa.

## 4.3 Výstupné dáta

Výsledkom našej metódy je rozpoznaná video sekvencia, pričom byť rozpoznaná znamená, že na vstupe graficky odlíšime rozpoznané postavy od ostatných objektov zachytených v scéne.

Činnosť metódy relevantne neovplyvní chod sekvencie v reálnom čase.

# Kapitola 5

## Implementácia

Aby bola naša metóda využiteľná v špecifickom prostredí podľa požiadaviek koncového užívateľa a čo najmenej obmedzovaná z našej strany, rozhodli sme sa metódu implementovať v podobe DirectShow filtra (sekcia 5.1.1), ktorý je možné zapojiť do rôznych DirectShow aplikácií, pričom je možné jeho chod stále upraviť nastavením jeho vlastností [16].

Od tohto rozhodnutia sa odvíja aj voľba programovacieho jazyka ktorým je C++ a vyvojové prostredie Microsoft Visual Studio 2003. Ďalšou výhodou prostredia C++ je možnosť pracovať s grafickou knižnicou OpenCV. Z nej sme okrem obrazových štruktúr využili aj implementáciu Kalmanovho a časticového filtra. [2]),

### 5.1 DirectShow

Microsoft DirectX SDK je nízko úrovňové aplikačné rozhranie (APIs) určené na vytváranie hier, alebo iných vysokovýkonných multimedialných aplikácií. Zahŕňa podporu 2D aj 3D grafiky, zvukových efektov, vstupných

zariadení v sieťových aplikáciách (napr. hry pre viac hráčov). DirectX SDK sa skladá z nasledovných celkov:

- DirectX Graphics
- Microsoft DirectInput
- Microsoft DirectPlay
- Microsoft DirectSound
- Microsoft DirectMusic
- Microsoft DirectShow
- DirectSetup
- DirectX Media

V našej práci využívame časť DirectShow, ktorá sa v poslednej dobe stala komponentom Windows Platform SDK.

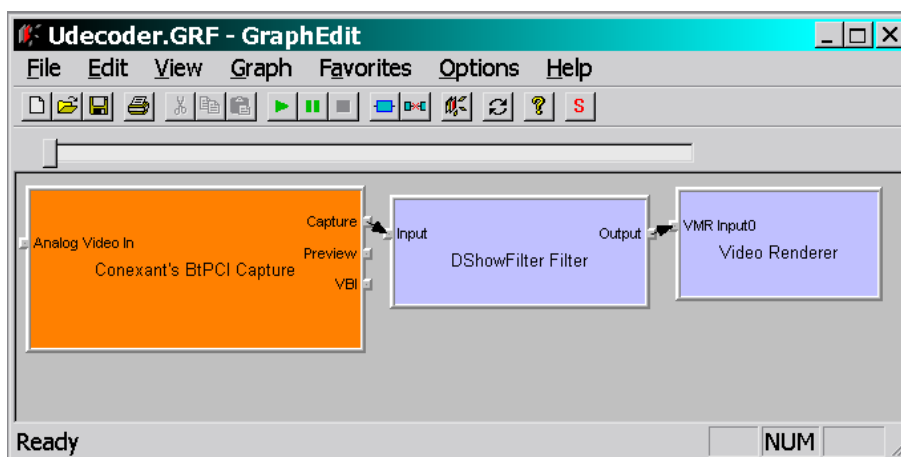
### **5.1.1 GraphEdit**

Aplikáciu poskytujúcu vizuálny ekvivalent funkčnosti SW, ktorý pracuje s DirectShow filtrami, nachádzame medzi základnými DirectShow nástrojmi a je to GraphEdit.

Umožňuje spájať rôzne filtre do grafu. Graf sa skladá zo samotných filtrov (vrcholov) pospájaných vstupnými a výstupnými pinmi (hranami), pričom ho je možné jednoducho editovať, takže užívateľovi poskytuje jednoduchú prácu s DirectShow.

## Filtre

Pri spájaní filtrov je potrebné dodržať pravidlá kladené na vstupné a výstupné piny (obr. 5.1).



Obr. 5.1: Filtre: Každý rám predstavuje filter. Šípky spájajúce filtre (vstup – výstup), a tiež vyjadrujú tok dát.

Rozpoznávame tri základné typy filtrov:

- Zdrojový (source filter)
- Transformačný (transform filter)
- Renderovací (render filter)

Zdrojový filter musí obsahovať každý graf, pretože slúži na získavanie zdrojových dát. Podporuje získavanie dát zo súboru, z URL, z TV-karty a tiež aj z iného zariadenia (napr. video záznam prostredníctvom firewire rozhrania).

Transformačný filter má za úlohu modifikovať ním prechádzajúci dátový tok, do jedného grafu je možné zapojiť niekoľko transformačných filtrov.

Poslednú časť z hľadiska smeru dátového toku tvorí rendrovací filter. Tento filter, reprezentuje napr. video okno – pre prehrávanie výstupného videa na obrazovku, zvuková karta – pre výstupný zvuk a súbor s ukladanými dátami, ktorý bude zapísaný na disk.

## Piny

Prostredníctvom pinov spájame jednotlivé filtre v grafe. Každý filter bez ohľadu na jeho typ, musí mať aspoň jeden pin pre spojenie s iným filtrom.

Pri spájaní dvoch filtrov, prechádzajú informácie ich pinmi a overí sa, či akceptujúci filter môže prijať dáta z odosielajúceho filtra. Ak sa piny úspešne rozpoznajú typ spracúvaných dát nastane spojenie medzi danými filtrami.

Na počet vstupných či výstupných pinov pre filter, nie sú kladené obmedzenia. Napr. MPEG-1 Stream Splitter filter potrebuje poslať audio a video časti oddelene dekodovacím filtrom.

## 5.2 Parametre nastaviteľné užívateľom

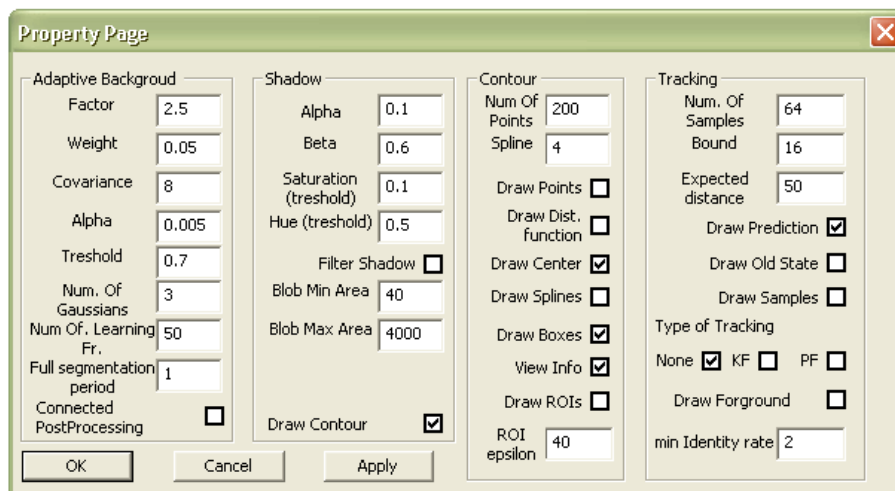
Aplikácia metód zameraných na počítačové videnie je vo väčšine prípadov závislá na špecifických vlastnostiach danej scény. Preto sme sa pri návrhu našej metódy snažili vytvoriť systém, ktorý by bol v prevažnej miere nastaviteľný užívateľom.

Na zmenu parametrov nám slúžia vlastnosti Direct Show filtra (property page), alebo dialóg Property page v Direct Show aplikácii (obr. 5.2).

Vlastnosťami filtra (aplikácie) môže užívateľ riadiť vizuálnu podobu výsledku a tiež parametre jednotlivých funkčných častí metódy, ktoré podrobne opíšeme v nasledovných sekciách <sup>1</sup>.

---

<sup>1</sup>Prednastavené hodnoty sú optimalizované na rozmer snímku sekvencie 320 × 240



Obr. 5.2: Dialóg vlastností (property page) systému

## 5.2.1 Adaptive Background

Factor

- reálne číslo
- prednastavená hodnota: 2, 5
- faktor pre určenie toho, či nový pixel vyhovuje niektorému z  $K$  Gaussiánov (kapitola 2).

Weight

- reálne číslo
- prednastavená hodnota: 0, 05
- počiatková váha Gaussiánu (kapitola 2).

Covariance



- reálne číslo
- prednastavená hodnota: 8
- kovariancia Gaussiánu (kapitola 2).

#### Alpha

- reálne číslo
- prednastavená hodnota: 0,005
- rýchlosť učenia – zmenšením tohto parametru spomalíme proces splývania nehybných objektov s pozadím. (kapitola 2)

#### Threshold

- reálne číslo
- prednastavená hodnota: 0,7 (kapitola 2)
- miera minimálneho podielu dát, ktoré môžu byť považované za pozadie, ak je hodnota tohto parametra vysoká, dostávame multimodálny vzor pozadia.

#### Num. Of Gaussians

- celé číslo
- prednastavená hodnota: 3
- určuje počet pravdepodobnostných (gaussovských) rozdelení s ktorými metóda adaptívneho pozadia pracuje,

- Každé jedno pravdepodobnostné rozdelenie v sebe obsahuje informáciu o farebnej pravdepodobnej hodnote obrazového bodu pozadia. Použitím väčšieho množstva pravdepodobnostných rozdelení, zvyšujeme schopnosť segmentácie, pracovať s pozadím, ktoré môže mať rôzne farebné hodnoty (napr. ak v scéne vlaje viacfarebná zástava),
- zmenou tohto parametra, sa systém nastaví na fázu učenia sa pozadia, pričom sa ostatné informácie o doteraz rozpoznávaných objektoch vynulujú.

#### Num. Of Learning Fr.

- celé číslo,
- prednastavená hodnota: 50
- určuje počet snímok vyhradený pre fázu učenia sa pozadia,
- zmenou tohto parametra, sa systém nastaví na fázu učenia sa pozadia, pričom sa ostatné informácie o doteraz rozpoznávaných objektoch vynulujú.
- Učenie by malo byť tak dlhé, aby sa počas neho pre daný obrazový bod scény, vystriedali všetky farebné hodnoty, ktoré pozadiu zodpovedajú.

#### Full segmentation period

- reálne číslo,
- prednastavená hodnota: 1
- určuje v akých intervaloch dochádza k úplnej segmentácii obrazového snímku (kapitola 2)

- zvýšením parametru urýchlíme chod systému, ale znížime počet snímkov v ktorých vyhodnocujeme objekty, ktoré na začiatku čiastočnej segmentácie v scéne neboli. Ak je interval príliš veľký v porovnaní s intenzitou svetelných zmien v scéne, môže sa stať, že pri nasledovnej úplnej segmentácii, segmentujeme ako pohyblivé objekty aj o časť pozadia.

## 5.2.2 Shadow

### Alpha

- reálne číslo
- prednastavená hodnota: 0, 1
- parameter dovoľuje vyhnúť sa identifikácii bodov ako tieňa, kde bolo pozadie ovplyvnené šumom (sekcia 2.2)

### Beta

- reálne číslo
- prednastavená hodnota: 0, 6
- parameter predstavuje „silu“ tieňa, čo je intenzita podľa podmienok prostredia (sekcia 2.2)

### Saturation (threshold)

- reálne číslo
- prednastavená hodnota: 0, 1
- prahovanie kanála S (sekcia 2.2)

Hue (threshold)

- reálne číslo
- prednastavená hodnota: 0,5
- prahovanie kanála H (sekcia 2.2)

Filter Shadow

- prepínač filtrovania tieňov
- prednastavený stav: vypnutý.

### 5.2.3 Blob

Blob Min Area

- celé číslo
- prednastavená hodnota: 40
- segmentované objekty, ktorých veľkosť (počet obrazových bodov) je menšia ako určuje parameter, nie sú systémom ďalej spracúvané.
- parameter, závisí od veľkosti relevantných objektov v scéne a veľkosti vstupnej sekvencie.

Blob Max Area

- celé číslo
- prednastavená hodnota: 40000
- segmentované objekty, ktorých veľkosť (počet obrazových bodov) je väčšia ako určuje parameter, nie sú systémom ďalej spracúvané.

- parameter, závisí od veľkosti relevantných objektov v scéne a veľkosti vstupnej sekvencie.

#### Blob Min Aspect Ratio

- reálne číslo
- prednastavená hodnota: 0,2
- segmentované objekty, ktorých pomer strán (aspect ratio) je menší ako určuje parameter, nie sú systémom ďalej spracúvané.
- parameter, závisí od tvaru ohraničujúceho obdĺžnika relevantných objektov v scéne.

#### Blob Max Aspect Ratio

- reálne číslo
- prednastavená hodnota: 1,0
- segmentované objekty, ktorých pomer strán (aspect ratio) je väčší ako určuje parameter, nie sú systémom ďalej spracúvané.
- parameter, závisí od tvaru ohraničujúceho obdĺžnika relevantných objektov v scéne.

#### Draw Contour

- prepínač zobrazenia všetkých obvodových bodov segmentovaného objektu,
- prednastavený stav: vypnutý.

## 5.2.4 Contour

Pre každý segmentovaný objekt pri rozpoznávaní spracúva iba výber z jeho bodov, ktorý označujeme ako kontúru (sekcia 6.4.4).

Num Of Points

- celé číslo
- prednastavená hodnota: 200
- určuje aký počet obvodových bodov segmentovaného objektu bude zaradený do kontúry (sekcia 6.4.4).

Spline

- celé číslo
- prednastavená hodnota: 4
- určuje v akých odstupoch<sup>2</sup> budú pre vzdialenostnú funkciu hľadané lokálne extrémny (sekcia 6.4.4).
- zväčšením parametru, odstránime šum na vzdialenostnej funkcii.

Draw Points

- prepínač zobrazenia kontúry obvodových bodov segmentovaného objektu
- prednastavený stav: vypnutý

Draw Dist. function

---

<sup>2</sup>step v rovnici (6.1)

- prepínač zobrazenia vzdialenostnej funkcie, v grafe (os  $x$ —body kontúry, os  $y$ —vzdialenosť od vzťažného bodu)
- prednastavený stav: vypnutý

#### Draw Center

- prepínač zobrazenia vzťažného bodu
- prednastavený stav: vypnutý

#### Draw Splines

- prepínač zobrazenia odstupov bodov kontúry, podľa ktorých sa hľadajú lokálne extrémny vzdialenostnej funkcie segmentovaného objektu
- prednastavený stav: vypnutý

#### Draw Boxes

- prepínač zobrazenia ohraničujúceho obdĺžnika segmentovaného objektu
- prednastavený stav: vypnutý

#### Draw ROIs

- prepínač zobrazenia ROI pre segmentované objekty (sekcia 6.2.2)
- prednastavený stav: vypnutý

#### View Info

- prepínač zobrazenia informácie o aktuálnom snímku sekvencie a počte segmentovaných oblastí, ktoré sú rozpoznávané
- prednastavený stav: vypnutý

## 5.2.5 Tracking

### Draw Prediction

- prepínač zobrazenia predpovedanej polohy ohraničujúceho obdĺžnika objektov v nasledujúcom snímku pri sledovaní objektov Kalmanovým alebo časticovým filtrom
- prednastavený stav: vypnutý

### Draw Old State

- prepínač zobrazenia polohy ohraničujúceho obdĺžnika objektov v predchádzajúcom snímku, pri sledovaní objektov Kalmanovým alebo časticovým filtrom
- prednastavený stav: vypnutý

### Draw Samples

- prepínač zobrazenia vzoriek pre predpoveď polohy ohraničujúceho obdĺžnika, pri sledovaní objektov časticovým filtrom
- segmentovaných oblastí, ktoré sú rozpoznávané
- prednastavený stav: vypnutý

### Type of Tracking

- prepínač spôsobu sledovania objektov pre čiastočnú segmentáciu (kapitola 6.2.2),
- prednastavený stav: None. (žiadne sledovanie)



# Kapitola 6

## Metóda na rozpoznávanie postáv (pedestrian tracker)

V tejto kapitole podrobne popíšeme činnosť metódy, ktorá dopĺňa iné relevantné práce z tejto vednej oblasti a vzájomné prepojenie jej jednotlivých funkčných modulov. Základným rozdielom našej metódy v porovnaní s prístupmi v sekcii 6.1 je rozpoznanie postavy na základe vlastností obrysu objektu ako vzdialenostnej funkcie.

Činnosť metódy je možné rozdeliť do troch základných modulov, ktoré sú medzi sebou prepojené. Tieto základné moduly sú segmentácia pozadia (kapitola 2) a, sledovanie objektu (kapitola 3) a rozpoznanie postavy (obr. 1.1), tieto operácie sú aplikované v každom snímku, okrem začiatkovej sekvencie, v ktorej prebieha len učenie – „inicializácia pozadia“.

### 6.1 Relevantné práce

V rozpoznávaní objektov – postáv sa stretávame s rôznymi prístupmi a zámermi. Výhody jednotlivých prístupov, sú závislé na vlastnostiach očakáva-

ného výsledku a počiatočnej situácie scény.

My sme sa rozhodli pre rozpoznávanie objektov z videosekvencie na základe vlastností obrysu. S podobným problémom, ale iným prístupom sa stretávame v niekoľkých zaujímavých prácach [8], [17], z ktorých najbližšie k našej metóde majú z hľadiska vstupných a výstupných hodnôt práce [3] a [18].

Práca [17] ponúka identifikáciu postáv za účelom odlišenia dvoch (viacerých) objektov, ktoré po segmentovaní tvoria súvislý celok, ale zachovávajú si farebnú odlišnosť

Identifikácia zameraná na 3D rekonštrukciu postavy je priblížená v štúdiu [8].

S identifikáciu podľa povahy obrysu daného objektu, pričom sú rozpoznávané tri triedy objektov: osoba, skupina osôb a motorové vozidlo iba na základe maximálnych lokálnych extrémov vzdialenostnej funkcie obrysu sa stretávame v článku [3].

Rozpoznanie postavy na základe naučenej siluety z 3D prostredia po snímaní ôsmimi virtuálnymi kamierami rozmiestnenými rovnomerne v okruhu tréningového 3D modelu postavy a po segmentovaní vstupnej videosekvencie metódou Canny-ho detekcie hrán spracúva systém opísaný v [18].

## 6.2 Segmentácia

Učenie pozadia je prvý krok nášho systému. Dĺžka učiacej sa sekvencie je závislá na rôznorodosti pozadia (svetelné zmeny, pohyb objektov z pozadia – stromy...) a počte pravdepodobných rozdelení s ktorými segmentovanie metódou adaptívneho pozadia ráta (kapitola 2).

Výsledkom segmentovania je pre náš systém binárna maska veľkosti vstup-

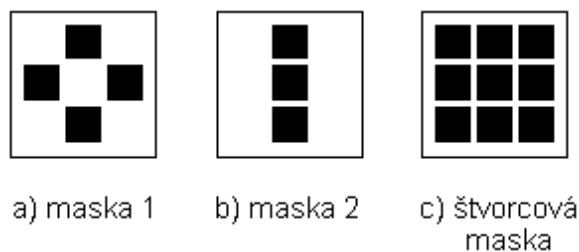
nej sekvencie. V oblastiach, ktoré predstavujú popredie, sa overí metódou opísanou v sekcii 2.2 či daný pixel masky reprezentuje tieň. Ak áno, priradíme mu hodnotu pozadia.

### 6.2.1 Dodatočné spracovanie

Masku získanú segmentáciou dodatočne spracujeme. Pre toto spracovanie sme sa rozhodli, aby sa dodatočne odstránili nežiadané artefakty, ktoré vznikajú segmentovaním vstupného obrazu, napr. v dôsledku šumu a kompresie videa.

S týmto krokom sa stretávame aj v iných metódach, kde použili napr. kombináciu: dvakrát dilatácia – jedenkrát erózia [3].

My sme zvolili dva alternatívne prístupy:



Obr. 6.1: Masky použité pre dodatočné spracovanie

**Prvá kombinácia:** erodovanie podľa vlastnej masky  $3 \times 3$  (obr. 6.1 a)), s prázdnyimi rohovými a stredovým bodom, dilatovanie podľa štvorcovej masky  $3 \times 3$  (obr. 6.1 c)) a opäť to isté erodovanie a po ňom tá istá dilatácia.

Pre tento postup sme sa rozhodli na základe priebežných výsledkov, počas ktorých sme spozorovali, že mnohé chybné segmentované oblasti mali charakter šumu, a ten sme sa týmto prístupom snažili zredukovať – rozdrobiť na

samostatné menšie oblasti.

**Druhá kombinácia:** otvorenie podľa štvorcovej masky  $3 \times 3$  (obr. 6.1 c)), otvorenie podľa vlastnej masky  $3 \times 3$  (obr. 6.1 b)), s prázdnyimi ľavými a pravými bodmi, dilatovanie podľa vlastnej masky  $3 \times 3$  (obr. 6.1 b)), dilatovanie podľa vlastnej masky  $3 \times 3$  (obr. 6.1 a)) s prázdnyimi rohovými a stredovým bodom.

Pre tento postup sme sa rozhodli, keď sme spozorovali, že ak sú pre nás relevantné objekty v snímku malé či nesúvislé prvou kombináciou ich poškodíme a nie je možné aby boli správne rozpoznané. Vtedy volíme opačný prístup ako v prvej kombinácii a to – oblasti čo najviac pospájať na súvislé väčšie oblasti.

Výber kombinácie je nastaviteľný zo strany užívateľa (sekcia 5.2).

## 6.2.2 Zrýchlenie

Keďže segmentovanie je proces, časovo závislý od veľkosti obrazu, rozhodli sme sa urýchliť túto metódu segmentovaním iba tej časti obrazu, na ktorej je možné očakávať objekt popredia. Z toho dôvodu sme zaviedli čiastočnú segmentáciu.

Segmentovať snímok čiastočne znamená, že metódu adaptívneho pozadia neaplikujeme na všetky obrazové pixle daného snímku, ale iba na vybrané pixle. Aby sme vedeli výber správne určiť pred čiastočnou segmentáciou, je nutné na predchádzajúci snímok aplikovať úplnú segmentáciu (vyhodnotiť metódu adaptívneho pozadia pre všetky obrazové body snímku) a vyhodnotiť z masky informácie o polohe objektov (presný postup sekcia 6.3).

Výber pixlov pre čiastočnú segmentáciu potom určujeme jedným z troch prístupov, ktoré je možné nastaviť zo strany používateľa (parametre) a to na

základe nasledujúcich možností:

- poloha ohraničujúceho obdĺžnika objektu,
- predikcia polohy ohraničujúceho obdĺžnika objektu Kalmanovým filtrom,
- predikcia polohy ohraničujúceho obdĺžnika objektu časticovým filtrom.

Týmto spôsobom dostávame pre každý objekt v obraze ohraničujúci obdĺžnik, ktorý je zväčšený o epsilónové okolie, tiež nastaviteľné zo strany užívateľa a pre každý objekt vymedzujeme ROI (Region of interest).

Výber pixlov vstupného snímku, na ktorých bude prebiehať čiastočná segmentácia, dostaneme zjednotením pixlov patriacich do jednotlivých ROI objektov.

Ak používame čiastočnú segmentáciu, dostávame sa do pozície, že objekty ktoré na scénu vstúpia, nie sú nájdené, pretože buď nie sú v častiach ktoré segmentujeme, alebo sú z časti obsiahnuté v ROI niektorých zo segmentovaných objektov. V tom prípade ich však opäť nenájdeme, pretože ak segmentujeme daný ROI objektu rozlíšime v ňom oblasť s najväčším obrysom – kontúrou a ten priradíme práve tomu objektu, ktorého ROI sme práve segmentovali.

### 6.3 Analýza masky – hľadanie objektov

Po segmentovaní získame masku, ktorú sme dodatočne spracovali morfológickými operáciami. Takto vzniknutá maska je binárna a obsahuje oblasti reprezentujúce pozadie a popredie, z ktorých je našim zámerom rozpoznať postavu človeka a elementárnym spôsobom vylúčiť čo najviac chybných výsledkov – objekty, ktorých vizuálny charakter je jasne odlišiteľný od postavy.

Z tohto dôvodu každú kontúru oblasti popredia, nájdenú v maske overujeme podľa kritérií na maximálnu a minimálnu veľkosť oblasti, ktorú ohraničuje a podľa maximálneho a minimálneho pomeru strán (aspect ratio) ohraničujúceho obdĺžnika danej kontúry. Kritéria sú nastaviteľné zo strany užívateľa, sekcia 5.2.

Ak nájdená kontúra spĺňa kritériá, je priradená novej oblasti (*blob-u*). Každá oblasť môže byť ďalej sledovaná Kalmanovým alebo časticovým filtrom. Podľa toho, pre aký typ čiastočnej segmentácie sa užívateľ rozhodol.

## 6.4 Rozpoznanie postavy

Táto časť je cieľom nášho systému. Pri návrhu štruktúry a funkčnosti našej metódy sme kládli dôraz na získavanie čo najväčšieho množstva informácie o objekte použitím čo najjednoduchších vyhodnotení. Takýmto prístupom však nie je možné dosiahnuť plnú robustnosť systému – úroveň rozpoznania ľudským okom a ľudským úsudkom.

Plná robustnosť je však nad rámec našej práce, pretože rôznorodosť reálnych vstupov je veľmi náročná a aplikácia by nepracovala v reálnom čase. Vo väčšine prípadov však systémy pracujú na vstupných dátach zo špecifických scén a preto ani plnú robustnosť nevyžadujú.

Aby sme sa k úplnej robustnosti čo najľahšie priblížili, je dobré si uvedomiť tri vlastnosti ľudského úsudku. Nevyhodnocuje scénu bez informácií o nej a ak áno, výsledok vypracúva zdĺhavejšie. Používa na rozpoznanie niekoľko prístupov podľa náročnosti danej reálnej situácie.

Ak uvážime, že každý človek alebo objekt má určitý tvar, určitú farbu a pohybové vlastnosti môžeme tieto prístupy rozdeliť na nasledovné:

- rozpoznanie podľa siluety – takto rozpoznávame napr. objekty pri sla-

bých svetelných podmienkach,

- rozpoznanie podľa farebných vlastností objektu, napr. na odlíšenie dvoch spoločne sa pohybujúcich postáv, či rozpoznanie bez pohybových informácií,
- rozpoznanie podľa pohybových vlastností objektu, najmä na rozpoznanie živých postáv od figurín.

V reálnom živote je nutné a pre človeka aj prirodzené tieto prístupy interaktívne kombinovať.

Pre „reálne“ pracujúcu aplikáciu je vhodné zvoliť jeden z prístupov, najvhodnejší pre vstupné situácie danej scény.

My sme sa rozhodli pre prvé kritérium – rozpoznanie podľa siluety. Z hľadiska tvaru je človek charakterizovaný svojou kostrou (rozmiestnením rúk, nôh, trupu a hlavy).

Ako prvé sme zvažovali dva prístupy:

- rozpoznanie postavy na základe kostry nájdenej pomocou morfologického algoritmu na hľadanie kostry ,
- rozpoznanie postavy pomocou algoritmu vzdialenostnej funkcie<sup>1</sup> (sekcia 6.4.2).

### 6.4.1 Kostra

Prednosťou tohto prístupu je, že dostaneme abstrahovanú informáciu o kostre objektu, ktorú môžeme priamo porovnať so zjednodušeným modelom kostry človeka.

---

<sup>1</sup>distance function

Tento algoritmus má veľkú časovú náročnosť, ktorá by priveľmi obmedzovala činnosť systému v reálnom čase.

### 6.4.2 Vzďialenostná funkcia

Týmto prístupom si pre postavu určíme funkciu, ktorá vyjadruje vzdialenosť obvodových bodov siluety, od vzťažného bodu (napr. ťažiska alebo stredu objektu). Túto funkciu použijeme ako vzor. Pre neznáme objekty určíme ich vlastné vzdialenostné funkcie, ktoré budeme so vzorom porovnávať. Nevýhoda spočíva vo veľkej početnosti vzorových vzdialenostných funkcií.

### 6.4.3 Kostra alebo vzdialenostná funkcia ?

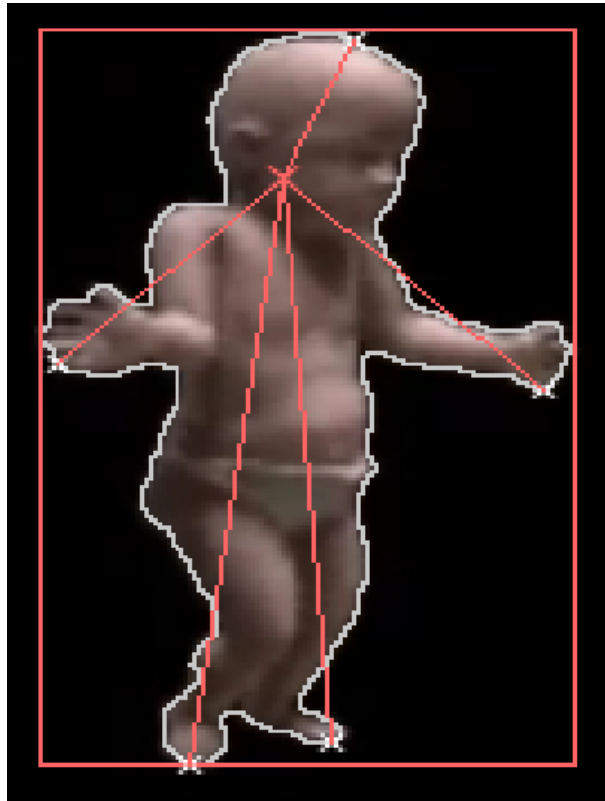
Obe z uvažovaných metód majú svoje kladné aj negatívne stránky. Preto sme sa rozhodli pre ich skĺbenie. Hľadali sme spôsob, akým by sme našli abstrahovanú informáciu o kostre zo vzdialenostnej funkcie pričom sme sa zamerali na pozitívne kritériá, čiže sme skúmali situácie v ktorých sme vyhodnocovali postavu človeka.

Pozorovali sme, že ak je postava vzpriamená a vzťažný bod pre výpočet vzdialenostnej funkcie je priemerom, všetkých bodov kontúry objektu, tak pre vzdialenostnú funkciu platí, že hlava, ruky aj nohy predstavujú lokálne maximá tejto funkcie.

Pospájaním lokálnych maxim spolu so vzťažným bodom sme dostali objekt podobný zjednodušenej kostre človeka (obr. 6.2).

Aby sme dosiahli silnejšie kritérium, rozhodli sme sa, že pre danú vzdialenostnú funkciu budeme vyhodnocovať aj lokálne minimá. Spájaním všetkých extrémov so vzťažným bodom vzdialenostnej funkcie síce nedostaneme abstrakciu kostry, ale spájaním extrémov vytvoríme jednoduchý model postavy (obr. 6.3).



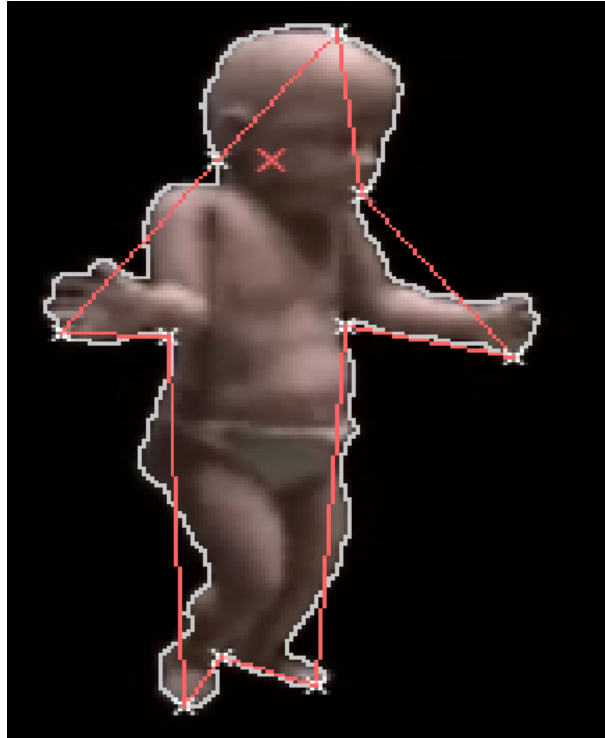


Obr. 6.2: Zjednodušený model kostry objektu, dosiahnutý pospájaním bodov kontúry objektu zodpovedajúcim maximálnym lokálnym extrémom vzdialenostnej funkcie s jej vzťažným bodom

S týmto prístupom, je problém rozpoznania postavy zredukovaný na problém nájdenia a overenia jednoduchého modelu postavy.

#### 6.4.4 Hľadanie extrémov

Na hľadanie kontúr z masky používame algoritmus pracujúci na báze 4-susednosti (ktorú sme pred 8-susednosťou uprednostnili z dôvodu rýchlejšieho vyhodnocovania). Pre urýchlenie metódy sme sa ďalej rozhodli, že nebudeme vzdialenostnú funkciu rátať na množine všetkých bodov kontúry.



Obr. 6.3: Zjednodušený model kontúry postavy, dosiahnutý pospájaním bodov kontúry objektu zodpovedajúcim lokálnych extrémov vzdialenostnej funkcie

Pri výbere bodov postupujeme nasledovne. Z každej nájdenej kontúry sme v danom snímku urobili pevný výber  $n$  bodov. Mohutnosť pevného výberu je nastaviteľná užívateľom (kapitola parametre). Body vyberáme z pôvodnej kontúry rovnomerne. Lokálne extrémny potom určujeme porovnaním troch po sebe idúcich bodov

$$x_i, x_{i+1}, x_{i+2}; 0 \leq i < n.$$

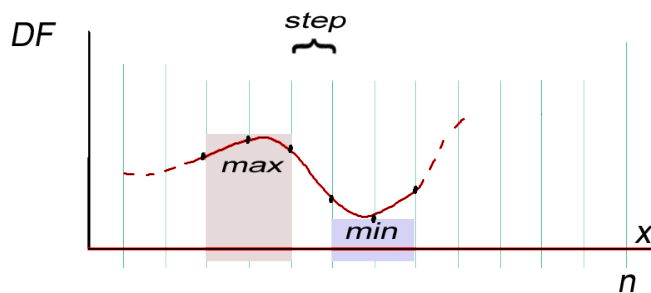
Keďže výber bodov obsahuje istý šum, rozhodli sme sa, že nebudeme porovnávať body

$$x_i, x_{i+1}, x_{i+2}; 0 \leq i < n,$$

ale

$$x_{i*\varepsilon}, x_{(i+1)*\varepsilon}, x_{(i+2)*\varepsilon}; 0 \leq i < \frac{n}{\varepsilon}. \quad (6.1)$$

$\varepsilon$  predstavuje parameter metódy (kapitola 5.2). Ak  $x_{(i+1)*\varepsilon}$  je maximom (minimom) spomedzi týchto bodov, hľadáme maximálnu (minimálnu) hodnotu vzdialenostnej funkcie na intervale  $x_{i*\varepsilon}, x_{(i+2)*\varepsilon}$  (obr. 6.4)



Obr. 6.4: Určovanie lokálnych extrémov pre vzdialenostnú funkciu

### 6.4.5 Vzťažný bod vzdialenostnej funkcie

V našej práci sme spočiatku klasicky používali ako vzťažný bod aritmetický priemer všetkých bodov kontúry. Neskôr sme sa rozhodli, že pre  $x$ -ovú súradnicu vzťažného bodu použijeme aritmetický priemer  $x$ -ových súradníc bodov kontúry.  $Y$ -ovú súradnicu vzťažného bodu sme určili ako 0,8 násobok výšky kontúry.

K tomuto rozhodnutiu nás viedlo pozorovanie, že ak má postava vystreté končatiny, a spojíme body zodpovedajúce lokálnym maximám vzdialenostnej funkcie s týmto vzťažným bodom úsečkou, zostane vnútri zjednoduše-

ného modelu postavy, ktorý získavame spájaním bodov kontúry, ktoré zodpovedajú lokálnym extrémom. Preto sa model vytvorený z takto nájdenej kontúry podobá na postavu viac, ako model vytvorený zo vzťažného bodu umiestneného v priemere bodov.

## 6.4.6 Rozpoznanie postavy

Pre každý neznámy objekt v scéne vyhodnotíme vzdialenostnú funkciu podľa nami navrhnutého určenia vzťažného bodu a určíme body kontúry, ktoré zodpovedajú lokálnym extrémom vzdialenostnej funkcie. Pozície týchto bodov sú rozhodujúce pri rozpoznávaní kontúry neznámeho objektu.

Základný zjednodušený model siluety postavy, určený lokálnymi extrémami vzdialenostnej funkcie, nezahŕňa prípady:

- postava má oblečenie vyčnievajúce z línie samotného tela postavy
- postava má v oblečení výraznú časť pliec (muži) alebo hrudne (ženy)

V reálnych situáciách sa ale s takýmito typmi postáv stretávame, v závislosti od klimatických podmienok či módných trendov. Preto sme sa rozhodli vytvoriť kategórie postáv podľa oblečenia nasledovne.

**Základ:** postava nemá oblečenie, ktoré by vyčnievalo z línie samotného tela postavy.

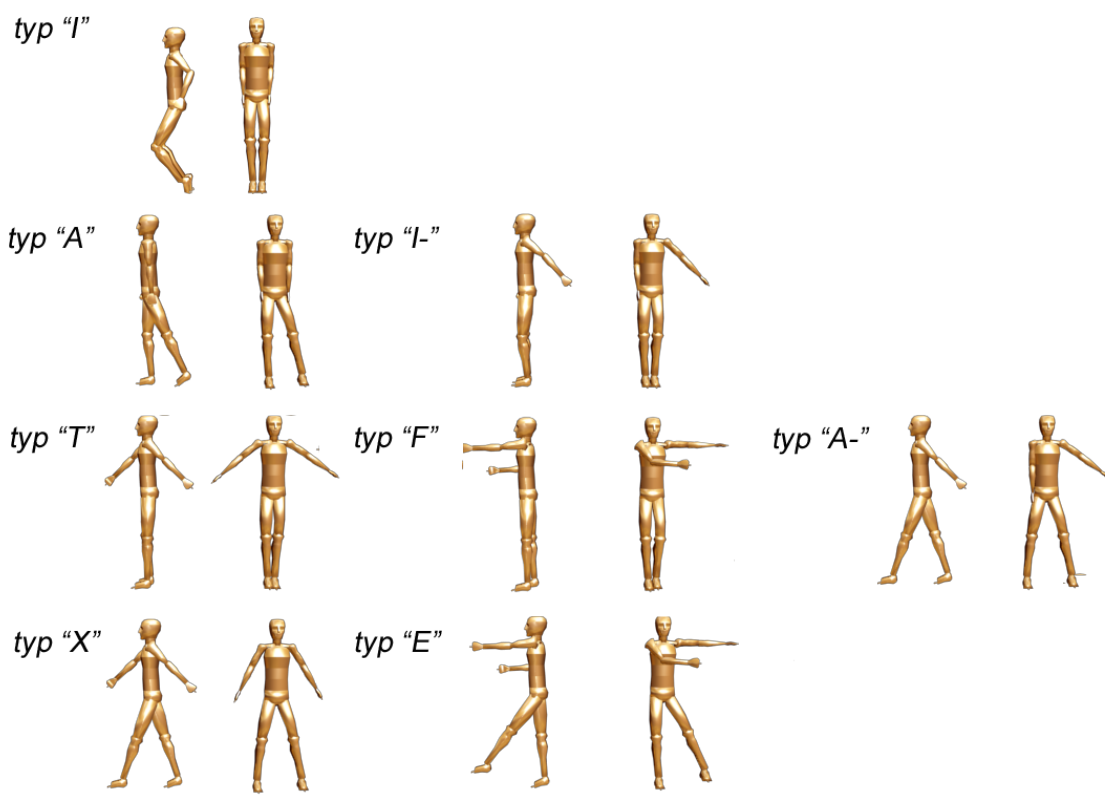
**Nohy:** nohavice vyčnievajú z línie samotných nôh (napr. postava má oblečené bermudy, širokú sukňu).

**Ruky:** rukávy vyčnievajú z línie horných končatín (napr. postava má oblečené tričko so širokými rukávami).

**Trup:** oblečenie vyčnieva z línie trupu (napr. postava má oblečenú voľnú vetrovku).

Určité narušenie línie spôsobené oblečením možno odstrániť pri hľadaní extrémov vzdialenostnej funkcie (sekcia 6.4.2). Preto sme s rozhodli implementovať kritéria na overenie extrémov iba pre základnú kategóriu.

Podľa počtu extrémov vzdialenostnej funkcie sme zaviedli nasledovné triedy póz pre základnú kategóriu postáv z pohľadu kamery (obr. 6.5), pričom sme neuvažovali prípady, ak má postava jednu z dolných končatín (alebo obe) nad úrovňou pásu alebo ak sú ruky vyššie ako hlava a z hľadiska siluety s ňou splývajú.



Obr. 6.5: Triedy póz pre základnú kategóriu postáv z pohľadu kamery

## Základná kategória

### 4 extrémny:

- „I“ – stojí znožmo, ruky má umiestnené, tak že splývajú s hrudnou časťou tela,

### 6 extrémov:

- „I“ – postava stojí znožmo, jednu ruku má umiestnenú, tak že splýva s hrudnou časťou tela, druhá vyčnieva do priestoru,
- „A“ – postava nestojí znožmo, ruky má umiestnené, tak že splývajú s hrudnou časťou tela,

### 8 extrémov:

- „T“ – postava stojí znožmo, ruky má umiestnené, tak že nesplývajú s hrudnou časťou tela ani s hlavou,
- „F“ – postava stojí znožmo, ruky má umiestnené, tak že nesplývajú s hrudnou časťou tela ani s hlavou a obe sú na jednej strane tela,
- „A“ – postava nestojí znožmo, jednu ruku má umiestnenú, tak že splýva s hrudnou časťou tela, druhá vyčnieva do priestoru,

### 10 extrémov:

- „X“ – postava nestojí znožmo, ruky má umiestnené, tak že nesplývajú s hrudnou časťou tela ani s hlavou,
- „E“ – postava nestojí znožmo, ruky má umiestnené, tak že nesplývajú s hrudnou časťou tela ani s hlavou a obe sú na jednej strane tela.

## 6.4.7 Rozpoznanie postavy

Objekt je rozpoznávaný ako postava ak:

- počet nájdených extrémov pre jeho vzdialenostnú funkciu nie je väčší ako 10
- hodnota každého nájdeného extrému patrí do prípustného intervalu hodnôt.

Prípustný interval hodnôt pre daný extrém určujeme osobitne podľa toho, aký typ pózy rozpoznávame<sup>2</sup> a či testujeme lokálne maximum (očakávame, že extrém reprezentuje: hlavu, dlaň či chodidlo) alebo lokálne minimum.

Lokálne minimum sa nachádzajú medzi lokálnymi maximami, preto sme ich rozlíšili na nasledovné typy:

- hlava – dlaň,
- hlava – chodidlo,
- dlaň – dlaň,
- chodidlo – chodidlo,
- dlaň – chodidlo.

Hodnoty vzdialenostnej funkcie v jednotlivých extrémoch, sme dali do pomeru s výškou testovanej kontúry aby sme dosiahli relatívnosť a prípustné hodnoty pre daný pomer sme určili na základe pozorovaní nasledovne:

- ak testujeme lokálne maximum:

– uvhlava : 0,5 – 0,3

---

<sup>2</sup>pripomínáme, že ten závisí na počte nájdených extrémov

- uvdlaň :  $0,2^3 - 0,4$
- uvchodidlo :  $0,2^4 - 0,85$
- ak testujeme lokálne minimum:
  - uvhlava – dlaň :  $0,0 - 0,29$
  - uvhlava – chodidlo:  $0,0 - 0,2^5$
  - uvdlaň – dlaň :  $0,2^6 - 0,4$
  - uvchodidlo – chodidlo:  $0,2^7 - 0,79$
  - uvdlaň – chodidlo :  $0,0 - 0,39$

Ak očakávame pózy typu „I“, „A“ testujeme či pomer strán ohraničujúceho obdĺžnika testovanej kontúry nieje nepomerne väčší k pomeru strán ohraničujúceho obdĺžnika daného zjednodušeného modelu.

### 6.4.8 Presnosť rozpoznania

Roztriedením póz postáv podľa počtu extrémov vzdialenostnej funkcie sme neobsiahli všetky reálne situácie. Je teda zrejmé, že v našom rozpoznávacom systéme budú negatívne vyhodnotené aj tie objekty, ktoré reprezentujú postavu. Preto je naša metóda do istej miery nepresná.

---

<sup>3</sup>minimálny pomer vzdialenostnej funkcie v bode mimimálnej šírky trupu človeka k výške kontúry

<sup>4</sup>minimálny pomer vzdialenostnej funkcie v najnižšom bode trupu človeka k výške kontúry

<sup>5</sup>minimálny pomer vzdialenostnej funkcie v bode maximálnej šírky trupu človeka k výške kontúry

<sup>6</sup>minimálny pomer vzdialenostnej funkcie v bode mimimálnej šírky trupu človeka k výške kontúry

<sup>7</sup>minimálny pomer vzdialenostnej funkcie v najnižšom bode trupu človeka k výške kontúry



Prípady ktoré uvažujeme sú pre správanie človeka v bežných situáciách prirodzené preto môžeme predpokladať, že postava pohybujúca sa v scéne, bude v niektorých snímkoch vstupnej sekvencie rozpoznaná pozitívne – správny výsledok. Analogicky sa môže stať, že objekt nereprezentujúci postavu, bude rozpoznán pozitívne – chybný výsledok.

Preto sa v našej práci opierame o predpoklad, že naše kritériá bude častejšie spĺňať objekt reprezentujúci postavu, ako iný objekt. Tento predpoklad reprezentuje parameter rozpoznania každého pohyblivého objektu v scéne. Užívateľ môže nastaviť prah ktorým vylúči objekty ktoré boli v scéne rozpoznané ako postava málokrát teda je nepravdepodobné, že ide o postavu.

Parameter rozpoznania sa určuje ako pomer počtu snímkov v ktorých bol objekt pozitívne rozpoznán (objekt je vyhodnotený ako postava) k počtu snímkov počas ktorých je objekt v scéne.

# Kapitola 7

## Testovacie dáta a výsledky

K aplikácii prikladáme aj testovacie súbory. Keďže naša práca je kompozíciou niekoľkých menších častí, vytvorili sme testovacie dáta zamerané na testovanie funkčnosti jednotlivých častí aj pre metódu ako celok.

Pre časti segmentácie a sledovania, sme zvolili syntetické dáta, aby sme overili ich základnú funkčnosť. Výsledky syntetických dát boli presné, čo potvrdzuje správnosť činnosti samostatných častí.

Pri testovaní na reálnych dátach sa stretávame najmä s nedostatočnými výsledkami segmentácie pozadia a filtráciou tieňov. Oba problémy sú z hľadiska robustnosti veľmi náročné, ale nie sú predmetom výskumu našej práce. Metódy ktoré sme použili sú závislé na nastavení vhodných parametrov, ktoré sú rôzne podľa charakteru scény (sekcia 5.2). Ako predvolené sme sa rozhodli použiť parametre odporúčané autormi [5] [11].

Pri výbere vstupných sekvencií z reálneho života sme sa zamerali na najčastejšie nastávajúce situácie: prechod chodcov a dopravných prostriedkov, s ohľadom na tri základné kritériá:

- pozadie scény

- svetelné podmienky
- poloha kamery

Naším cieľom bolo pripraviť aj také videá, na ktorých sa ukážu hranice našej metódy, preto nie všetok testovací materiál spĺňa požiadavky danej špecifikáciou.

Testovacie videá sú komprimované. Pre uniformnosť som použila to isté kódovanie, ktoré vyhovovalo našim technickým možnostiam. Pre úplnosť sme k testovacím dátam pridali aj kodek (ffcodec div3).

### 7.0.9 Testovací stroj

AMD Athlon(tm) 64 2800+ (1,8 GHz), 768 MB of RAM.

## 7.1 Rýchlosť

Pri testovaní videosekvencií sme použili zväčša predvolené hodnoty parametrov.

Test rýchlosti chodu metódy v závislosti od obrazového rozlíšenia a veľkosti periódy úplnej segmentácie sme vykonali pre testovacie videosekvencie odlišné iba vo veľkosti snímku. Pre typ čiastočnej segmentácie sme volili nastavenie ROI podľa ohraničujúceho obdĺžniku. Získané výsledky sú v priemernom<sup>1</sup> prípade nasledovné:

(320x240) / (640x480)

perióda úplnej segmentácie: 1 – počet snímok za sekundu: 19 fps / 5 fps

---

<sup>1</sup>napr. S\_50fr\_postava(320 × 240).avi

perióda úplnej segmentácie: 10 – počet snímkov za sekundu: 29 fps / 10 fps

perióda úplnej segmentácie: 20 – počet snímkov za sekundu: 30 fps / 10 fps

Test rýchlosti chodu metódy v závislosti od typu segmentácie (úplnej alebo čiastočnej (sekcia 6.2.2) ) a voľby typu sledovania nasledovné:

rozlíšenie:  $640 \times 480$ , perióda úplnej segmentácie: 1

sledovanie: žiadne – počet snímkov za sekundu: 5fps

sledovanie: Kalmanov filter – nepoužíva sa

sledovanie: časticový filter – nepoužíva sa

rozlíšenie:  $640 \times 480$ , perióda úplnej segmentácie: 10

sledovanie: žiadne – počet snímkov za sekundu: 10 fps

sledovanie: Kalmanov filter – počet snímkov za sekundu: 12,4 fps

sledovanie: časticový filter – počet snímkov za sekundu: 12 fps

rozlíšenie:  $640 \times 480$ , perióda úplnej segmentácie: 20

sledovanie: žiadne – počet snímkov za sekundu: 10fps

sledovanie: Kalmanov filter – počet snímkov za sekundu: 14,8 fps

sledovanie: časticový filter – počet snímkov za sekundu: 13 fps

## 7.2 Presnosť

Pre scény v ktorých sa objekt vyskytuje dlhšie na jednom mieste, sme pre segmentačný parameter alpha zvolili hodnotu 0,0005. Pre dosiahnutie lepších výsledkov sme menili typ dodatočného spracovania (postprocessing), počet

výberových bodov pre kontúru či odstup pre hľadanie extrémov (spline) podľa veľkosti očakávanej postavy.

Táto sekcia je zameraná na presnosť výsledkov<sup>2</sup>, ktoré treba overovať manuálne<sup>3</sup>.

### 7.2.1 Exteriér

Hodnoty uvádzame v poradí: pozitívne vyhodnotené testy, všetky vyhodnotené testy, pomer pozitívnych ku všetkým testom.

objekt je postava: 17, 21, 76.

objekt nieje postava: 20, 6, 100<sup>4</sup>.

### 7.2.2 Interiér

Hodnoty uvádzame v poradí: pozitívne vyhodnotené testy, všetky vyhodnotené testy, pomer pozitívnych ku všetkým testom.

objekt je postava: 19, 32, 59

objekt nieje postava: 1, 9, 11

### 7.2.3 Syntetické scény

Hodnoty uvádzame v poradí: pozitívne vyhodnotené testy, všetky vyhodnotené testy, pomer pozitívnych ku všetkým testom.

objekt je postava: 186, 517, 35

objekt nieje postava: 2, 36, 5

---

<sup>2</sup>Automatické overovanie výsledkov nie je možné, pretože to vedie k „zacykleniu“ :-)

<sup>3</sup>Preto pre presnejšiu predstavu o výsledkoch, odporúčame testovať priložené testovacie dáta.

<sup>4</sup>vo väčšine prípadov v reálnej scéne objekt ako postavu zamietneme na základe veľkosti a pomeru strán

### 7.3 Výsledky iných

Porovnanie našich výsledkov s výsledkami kolegov, v tejto oblasti veľmi náročné, pretože vo veľkej miere závisí od daného vstupu, ktorý však pre nás nieje prístupný. Aby sme my umožnili čo najrôznorodjšie porovnanie, poskytli sme niektoré z našich testovacích dát na CD nosiči priloženom k textu. Z hľadiska rýchlosti je porovnanie zavádzajúce, pretože netestujeme na rovnako silných strojoch, pričom treba pripomenúť, že všetky metódy sa snažia (často úspešne) o chod v reálnom čase.

Detailné opísanie výsledkov iných prác je nad rámec našej témy, preto odporúčame práce [3] a [18].

# Kapitola 8

## Budúca práca

Keďže počítačové videnie je dynamicky rozvíjajúca sa oblasť výskumu radi by sme náš systém v budúcnosti zdokonalili. Vylepšenia by sme chceli uskutočniť najmä v zrýchlení systému a v presnejších výsledkoch.

Uvažujeme o implementovaní nášho systému v jazyku nižšej úrovne a v implementácii na grafickej karte.

Pre dosiahnutie lepších výsledkov, by sme radi implementovali robustnejšiu metódu pre filtrovanie tieňov a pridali by sme triedy rozpoznania aj pre postavy v iných ako vzpriamených polohách, ktoré sme do systému nezaradili v tejto verzii, pretože sme boli limitovaný rýchlosťou rozpoznania, ktorú požadujeme v reálnom čase.

# Kapitola 9

## Záver

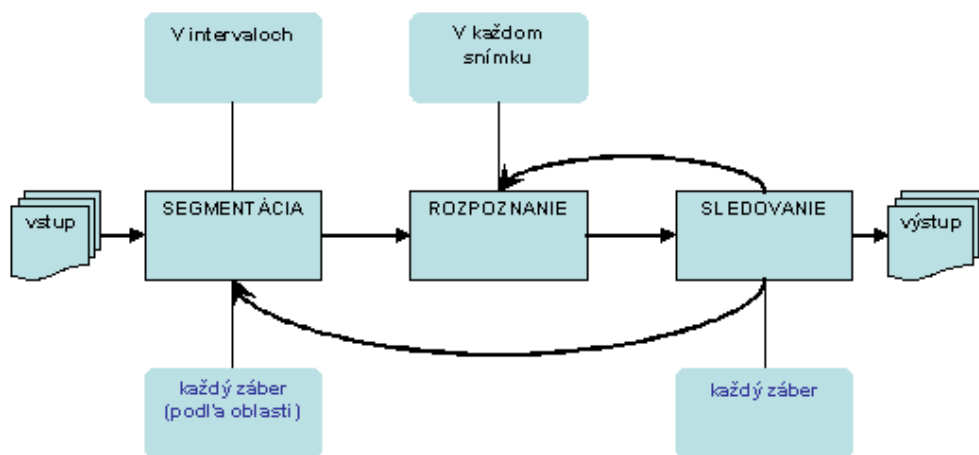
V našej práci sme vytvorili systém určený pre rozpoznávanie postáv v sekvencii obrazov. Jeho činnosť stručne charakterizuje obr. 9.1. Pre segmentovanie sme zvolili metódu adaptívneho pozadia, pre sledovanie Kalmanov a časticový filter.

Pre detekciu postavy sme implementovali vlastnú metódu, založenú na zjednodušení tvaru obrysu rozpoznávaného objektu. Tvar obrysu analyzujeme prostredníctvom vzdialenostnej funkcie, ktorá je v každom obrazovom snímku vyhodnocovaná pre segmentované objekty. Pre  $x$ -ovú súradnicu vzťažného bodu používame aritmetický priemer  $x$ -ových súradníc bodov kontúry.  $Y$ -ovú súradnicu vzťažného bodu sme určili ako 0,8 násobok výšky kontúry.

Po nájdení lokálnych extrémov vytvoríme zjednodušený model kontúry, čím dosahujeme rýchlu a zároveň výstižnú informáciu o povahe siluety objektu. Rýchlosť vytvorenia zjednodušeného modelu siluety nám umožňuje vykonávať rozpoznanie v každom snímku a tým pre každý objekt vyjadriť pravdepodobnosť toho, či je rozpoznávaný objekt skutočne postavou. Ak je pomer pozitívne vyhodnotených testov rozpoznania k pomeru všetkých vyko-



naných testov pre daný objekt väčší ako užívateľom zvolená hodnota, objekt je vizuálne spracovaný ako postava.



Obr. 9.1:

# Literatúra

- [1] Dengsheng Zhang and Guojun Lu. Segmentation of moving objects in image sequence: A review.
- [2] Intel Corporation. Open source computer vision library - reference manual. 2001.
- [3] Murat EKINCI and Eyup GEDIKL.
- [4] Michal Jůza. Detection of the unusual behavior from video sequences. 2004.
- [5] Chris Stauffer and W.E.L Grimson. Adaptive background mixture models for real-time tracking.
- [6] P. Wayne Power and Johann A. Schoonees. Understanding background mixture models for foreground segmentation. 2002.
- [7] Thanarat Horprasert, David Harwood, and Larry S. Davis. A statistical approach for real-time robust background subtraction and shadow detection.
- [8] Ronald Poppe, Dirk Heylen, Anton Nijholt, and Mannes Poel. Towards real-time body pose estimation for presenters in meeting environments. 2005.

- [9] K. Wojciechowski, B. Smolka, H. Palus, R.S. Kozera, W. Skarbek, and L. Noakes. Competervision and graphics.
- [10] Andrea Prati, Ivana Miki, Costantino Grana, and Mohan M. Trivedi. Shadow detection algorithms for traffic flow analysis: a comparative study.
- [11] Rita Cucchiara, Costantino Grana, Massimo Piccardi, Andrea Prati, and Stefano Sirotti. Improving shadow suppression in moving object detection with hsv color information.
- [12] Greg Welch and Gary Bishop. An introduction to the kalman filter. 2004.
- [13] Ľubomír Molent. Semestrálny projekt.
- [14] Mun Wai Lee, Isaac Cohen, and Soon Ki Jung. Particle filter with analytical inference for human body tracking.
- [15] Michael Isard and Andrew Blake. Condensation - conditional density propagation for visual tracking. 1998.
- [16] DirectX 9.0 programmer's reference. 2002.
- [17] Sangho Park and J.K. Aggarwal. Simultaneous tracking of multiple body parts of interacting persons.
- [18] M. Dimitrijevic, V. Lepetit, and P. Fua. Human body pose recognition using spatio-temporal templates.